SCOTTISH
QUALIFICATIONS
AUTHORITY

2005 Applied Mathematics

Advanced Higher – Statistics

Finalised Marking Instructions

These Marking Instructions have been prepared by Examination Teams for use by SQA Appointed Markers when marking External Course Assessments.

General Marking Principles

These principles describe the approach taken when marking Advanced Higher Mathematics papers. For more detailed guidance please refer to the detailed Marking Instructions.

1 The main principle is to give credit for the skills demonstrated and the criteria met. Failure to have a correct method may not preclude a candidate gaining credit for their solution.

2 The answer to one part of a question, even if incorrect, is accepted as a basis for subsequent dependent parts of the question.

3 The following are not penalised:

- working subsequent to a correct answer (unless it provides firm evidence that the requirements of the question have not been met)
- legitimate variation in numerical values / algebraic expressions.

4 Full credit will only be given where the solution contains appropriate working. Where the correct answer might be obtained by inspection or mentally, credit may be given.

5 Sometimes the method to be used in a particular question is explicitly stated; no credit will be given where a candidate obtains the correct answer by an alternative method.

6 Where the method to be used in a particular question is not explicitly stated, full credit will be given for an alternative valid method.

In the detailed Marking Instructions which follow, marks are shown alongside the line for which they are awarded. There are two codes used, M and E. M indicates a method mark, so in question 1, 1M, 1, 1 means a method mark for the product rule (and then a mark for each of the terms). E is shorthand for error. In question 3, 2E1, means that a correct answer is awarded 2 marks but that 1 mark is deducted for each error.

Section A – Statistics

A1. For recognising that the required probability is $P(\bar{S} \mid C)$ **1**

$$= \frac{P(\bar{S} \cap C)}{P(C)}$$

$$= \frac{P(C \cap \bar{S})}{P(C)}$$

$$= \frac{P(C \mid \bar{S}) . P(\bar{S})}{P(C \mid \bar{S}) . P(\bar{S}) + P(C \mid S) . P(S)} \qquad \mathbf{1}$$

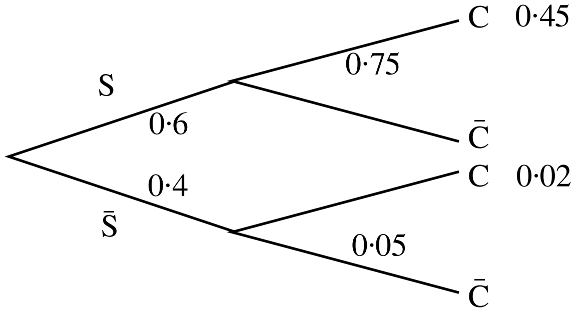
$$= \frac{0{\cdot}05 \times 0{\cdot}4}{0{\cdot}05 \times 0{\cdot}4 + 0{\cdot}75 \times 0{\cdot}6} \qquad \mathbf{1, 1}$$

$$= \frac{0{\cdot}02}{0{\cdot}02 + 0{\cdot}45}$$

$$= 0{\cdot}043\ldots \text{ or } \frac{2}{47} \qquad \mathbf{1}$$

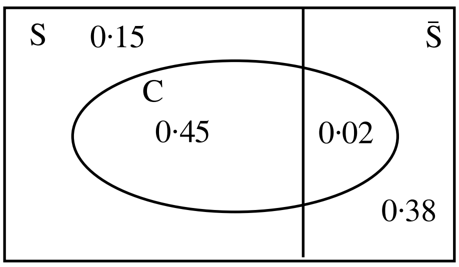
Alternative methods:

1. *Tree diagram*:



This diagram (fully labelled) gets the second and third marks and can then lead to the final 2 marks.

2. *Venn diagram*:



This diagram (with correct probabilities included) gets the second and third marks and can then lead to the final 2 marks.

A2. (a) By symmetry the mean $\mu = 22$. **1**

$$\frac{26 - 22}{\sigma} = 1{\cdot}96 \qquad \textbf{1, 1}\ (\dagger)$$

and the standard deviation $\sigma = 2{\cdot}04$. **1**

(†) If 1·64 is used instead of 1·96, deduct 1 mark.
Corresponding $\sigma = 2{\cdot}44$ gains the third and fourth marks.

(b)
$$\begin{aligned} \mathrm{P}(\mathrm{BMI} > 28) &= 1 - \mathrm{P}(\mathrm{BMI} \leqslant 28) && \textbf{1} \\ &= 1 - \mathrm{P}\left(Z \leqslant \frac{28 - 22}{2{\cdot}04}\right) \\ &= 1 - \mathrm{P}(Z \leqslant 2.94) && \textbf{1} \\ &\approx 0{\cdot}0016 \text{ (or } 0{\cdot}002). && \textbf{1} \end{aligned}$$

Note: The value 1·64 in part (a) leads to 0·0069 which would get all 3 marks in part (b).

A3. (a)
$$\hat{p} = \frac{531}{1000} = 0{\cdot}531$$
$$\hat{q} = 1 - \hat{p} = 0{\cdot}469. \qquad \textbf{1}$$

The required confidence interval is

$$\begin{aligned} &\hat{p} \pm 2{\cdot}58\sqrt{\frac{\hat{p}\hat{q}}{n}} && \textbf{1} \\ &= 0{\cdot}531 \pm 2{\cdot}58\sqrt{\frac{0{\cdot}531 \times 0{\cdot}469}{1000}} \\ &= 0{\cdot}531 \pm 0{\cdot}041 \text{ or } (0{\cdot}490,\ 0{\cdot}572) && \textbf{1} \end{aligned}$$

Since the confidence interval includes 0.5 one cannot reject the hypothesis that half of adults snack regularly on chocolate. **1**

(b)
- No sampling frame is available for the population so random sampling methods cannot be used.
- Respondents may be reluctant to be truthful about their snacking habits.
- The amount of time needed is likely to be an issue.

Any of these can earn a mark. **1,1**

(c) Acceptable:
Advantages:
Telephone surveys are relatively cheap to conduct; or,
quick to do since there is no travel; **1**
Disadvantages:
Using only respondents with access to a telephone is a potential source of bias since not everyone has a phone. **1**

[Unacceptable:
Better response rate by phone; more likely to talk to you than in the street; not everyone has a phone; not everyone might be in when you phone.]

A4. Assumptions are that the sample is random (not 'independent') (not 'still normal') **1**
and the standard deviation is unchanged. **1**
The mean of the sample is $\bar{x} = 26{\cdot}7$. **1**
$H_0 : \mu = 29$
$H_1 : \mu < 29$. **1**

The critical region for a z-test at the 5% level of significance is $z < -1{\cdot}64$. **1**

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{26{\cdot}7 - 29}{3/\sqrt{10}} = -2.42.$$ **1**

Since $-2{\cdot}42 < -1{\cdot}64$, the null hypothesis that the mean is 29 would be rejected and so the data provide evidence of reduced mean cycle time. **1**

- This might be answered by using a CI but would lose a mark for using a 2-sided interval. The z-interval works out as (24.84, 28.56) so as this does not contain the value 29, H_0 should be rejected.
- A t-interval with $s = 3{\cdot}33$, $(24{\cdot}31, \ 29{\cdot}09)$ might be used but it has to be followed by not rejecting H_0. (This would gain a maximum of 5 marks.)

A5.

$$s^2 = \frac{\text{SSR}}{n - 2}$$ **1**

$$= \frac{0{\cdot}12454}{10} \qquad \therefore s = 0{\cdot}11160.$$ **1**

$$t_{10,\,0{\cdot}975} = 2{\cdot}228$$ **1**

The required confidence interval is given by

$$-1{\cdot}138 \pm 2{\cdot}228 \times \frac{0{\cdot}11160}{\sqrt{2{\cdot}58309}}$$ **1**

$$= -1{\cdot}138 \pm 0{\cdot}155 \text{ or } (-1{\cdot}293, \ -0{\cdot}983)$$ **1**

Since the confidence interval includes -1, Zipf's law is supported by the result. **1**
[Must make reference to 'includes -1' for this mark.]

A6. (a) Let X denote the number of correct answers obtained.
$X \sim \text{Bin}(12, \ 0{\cdot}25)$ **1**

$$\text{P(pass)} = \text{P}(X \geqslant 8)$$

$$= 1 - \text{P}(X \leqslant 7)$$ **1**

$$\approx 1 - 0{\cdot}9972$$

$$= 0{\cdot}0028.$$ **1**

(b) $X \sim \text{Bin}(12, \ 0{\cdot}5)$ [or just $p = 0{\cdot}5$] **1**

$$\text{P(pass)} \approx 1 - 0{\cdot}8062 = 0{\cdot}1938.$$ **1**

A7. (a)

$$\bar{x} = \frac{\text{Total no. of impacts}}{\text{Total no. of squares}}$$

$$= \frac{229 \times 0 + 211 \times 1 + 93 \times 2 + 35 \times 3 + 7 \times 4 + 1 \times 5}{576}$$

$$= \frac{535}{576} = 0{\cdot}93 \text{ (or other valid approximation)}$$ **1**

The Poi (0·93) distribution has $P(X = x) = f(x) = \dfrac{e^{-0{\cdot}93}0{\cdot}93^x}{x!}$ **1**

n	Expected frequency of squares with n impacts	
1	$576 \times f(1) = 576 \times e^{-0.93}\dfrac{0{\cdot}93}{1!} = 211{\cdot}4$	**1,1** (†)
2	$576 \times f(2) = 576 \times e^{-0.93}\dfrac{0{\cdot}93^2}{2!} = 98{\cdot}3$	
3	$576 \times f(3) = 576 \times e^{-0.93}\dfrac{0{\cdot}93^3}{3!} = 30{\cdot}5$	**1**

(†) The first of these two marks is for the idea of $576 \times f(1)$.

(b)

Impacts i	No. of squares	Expected no. of squares		
	O_i	E_i	$(O_i - E_i)^2$	$\dfrac{(O_i - E_i)^2}{E_i}$
0	229	227·3	2·89	0·01
1	211	211·4	0·16	0·00
2	93	98·3	28·09	0·29
3	35	30·5	20·25	0·66
4	7	7·1	0·01	0·00
5 or more	1	1·4	0·16	0·11
			Total	1·07

Last 2 cells combined to regulate expected frequencies. **1**(†)

H_0: Observations are from the Poi(0.93) distribution.

H_1: Observations are not from the Poi(0.93) distribution. **1**

Chi-squared test-statistic ≈ 1·07 **1**

Degrees of freedom $6 - 1 - 1 = 4$

Critical value of Chi-squared at 5% level of significance = 9·488 **1**

Since 1·07 < 9·488 the null hypothesis that the number of impacts per square has a Poisson(0.93) distribution is accepted. **1,1**

(†) This mark would be lost if expected frequencies were dealt with in a different way.

(c) Since a Poisson distribution implies randomness. **1**

A8. (a)

M1		M2
9	5	
2 9 6 4 8	6	9 6
2 7 0	7	
	8	1 6
	9	0 3 2

| 5 | 1 Represents 5.1%Ag **2**

This suggests that M2 has a higher silver content. **1**

Other acceptable displays include:

A box-plot on a common scale or

some other visualisation which clearly shows the comparison.

A simple line of ranked data is not enough (and cannot gain the first 2 marks.)

(b) Rank sum for Second Mintage

$$W = 4{\cdot}5 + 7{\cdot}5 + 12 + 13 + 14 + 15 + 16 = 82$$ **1**

$H_0 : \eta_1 = \eta_2$

$H_1 : \eta_1 \neq \eta_2$ **1**

$$\mathrm{E}(W) = \tfrac{1}{2}n(n + m + 1) = 59{\cdot}5$$

$$\mathrm{V}(W) = \frac{1}{12}nm(n + m + 1) = 89{\cdot}25 = 9{\cdot}45^2$$ **1** (†)

$$z = \frac{W - \mathrm{E}(W)}{\sqrt{\mathrm{V}(W)}} = \frac{82 - \frac{1}{2} - 59{\cdot}5}{9{\cdot}45} = 2{\cdot}33$$ **1** (†)

The critical region for the 5% significance level is $|z| > 1{\cdot}96$. **1**

Since 2.33 lies in the critical region the null hypothesis is rejected **1**

and so there is evidence that the median silver content differs for the two mintages. **1**

(†) These marks cannot be awarded if a normal approximation is not used.

A9. (a) The 3 sigma limits are $\mu \pm 3\frac{\sigma}{\sqrt{n}}$ **1**

i.e. $215 \pm 3\frac{0{\cdot}1}{\sqrt{4}} = 215 \pm 0{\cdot}15$ or $(214{\cdot}85,\ 215{\cdot}15)$. **1**

(b) None of the plotted points lie above the upper 3 sigma limit or below the lower 3 sigma limit so there is no evidence of any problem and the manager should be satisfied that the process is operating in a state of statistical control. **1**

(c)

$$\begin{aligned} \text{P(Point outside 3 sigma limits)} &= 1 - \text{P}\left(\mu - 3\frac{\sigma}{\sqrt{n}} \leqslant \overline{X} \leqslant \mu + 3\frac{\sigma}{\sqrt{n}}\right) && \mathbf{1} \\ &= 1 - \text{P}(-3 \leqslant Z \leqslant 3) && \mathbf{1} \\ &\approx 2(1 - 0{\cdot}9987) \\ &= 0{\cdot}0026 && \mathbf{1} \end{aligned}$$

[For this, a value of 0·1336 should be given 1 (out of 3)].

$$\begin{aligned} \text{P(Point below 2 sigma lower limit)} &= \text{P}(Z < -2) \\ &= 1 - \text{P}(Z \leqslant 2) \\ &\approx 1 - 0{\cdot}9772 \\ &= 0{\cdot}0228 && \mathbf{1} \end{aligned}$$

[A value of 0·1587 should be given 1].

This probability is less than 0.05 and so in terms of hypothesis testing it could be regarded as significant and therefore of concern to the manager as possible evidence of a special cause affecting the process. **1**

[To obtain this mark, the concern should relate either to the calculated probability or to the warning limits (with opposite conclusions).]

Using warning limits (alone) can only gain the final mark i.e. 1 out of 5.

Section B – Mathematics

B1. (a)

$$f(x) = \exp\left(\tan \tfrac{1}{2}x\right)$$

$$f'(x) = \sec^2 \tfrac{1}{2}x \left(\tfrac{1}{2}\right) \exp\left(\tan \tfrac{1}{2}x\right) \qquad \textbf{1,1,1}$$

$$= \tfrac{1}{2} \sec^2 \tfrac{1}{2}x \exp\left(\tan \tfrac{1}{2}x\right)$$

(b)

$$g(x) = \left(x^3 + 1\right) \ln\left(x^3 + 1\right)$$

$$g'(x) = 3x^2 \ln\left(x^3 + 1\right) + \left(x^3 + 1\right)\frac{3x^2}{x^3 + 1} \qquad \textbf{1,1}$$

$$= 3x^2 \ln\left(x^3 + 1\right) + 3x^2 \qquad \textbf{1}$$

B2.

$$A^2 - A = \begin{pmatrix} 2 & 1 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 0 & -1 \end{pmatrix} - \begin{pmatrix} 2 & 1 \\ 0 & -1 \end{pmatrix} \qquad \textbf{M1}$$

$$= \begin{pmatrix} 4 & 1 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 2 & 1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \qquad \textbf{1}$$

$$= 2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \textbf{1}$$

B3.

$$x = 0 \Leftrightarrow 5t^2 - 5 = 0 \Leftrightarrow t = \pm 1 \qquad \textbf{M1}$$

$$y = -3 \Leftrightarrow 3t^3 = -3 \Leftrightarrow t = -1$$

At $(0, -3)$, $t = -1$. **1**

$$\frac{dx}{dt} = 10t; \ \frac{dy}{dt} = 9t^2; \qquad \textbf{1}$$

$$\frac{dy}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}} = \frac{9t}{10} \qquad \textbf{1}$$

So when $t = -1$, the gradient is $-\dfrac{9}{10}$. **1**

B4.

$$\left(2a - \frac{3}{a}\right)^4 = (2a)^4 + 4(2a)^3\left(-\frac{3}{a}\right) + 6(2a)^2\left(-\frac{3}{a}\right)^2 + 4(2a)\left(-\frac{3}{a}\right)^3 + \left(-\frac{3}{a}\right)^4 \qquad \textbf{1 powers} \quad \textbf{1 coeff}$$

$$= 16a^4 - 96a^2 + 216 - \frac{216}{a^2} + \frac{81}{a^4} \qquad \textbf{1}$$

B5.

$$\frac{x^2 + 3}{x\left(1 + x^2\right)} = \frac{A}{x} + \frac{Bx + C}{1 + x^2} \qquad \textbf{M1}$$

$$x^2 + 3 = A\left(1 + x^2\right) + (Bx + C)x$$

$$x = 0 \quad \Rightarrow \quad 3 = A \qquad \textbf{1}$$

$$x^2 + 3 = 3\left(1 + x^2\right) + (Bx + C)x$$

$$x = 1 \quad \Rightarrow \quad 4 = 6 + B + C$$

$$x = -1 \quad \Rightarrow \quad 4 = 6 + B - C$$

$$2C = 0 \Rightarrow C = 0 \text{ and } B = -2 \qquad \textbf{1}$$

$$\int_{1/2}^{1} \frac{x^2 + 3}{x\left(1 + x^2\right)}\, dx = \int_{1/2}^{1} \frac{3}{x} - \frac{2x}{1 + x^2}\, dx$$

$$= \left[3 \ln x - \ln\left(1 + x^2\right)\right]_{1/2}^{1} \qquad \textbf{1}$$

$$= \left[0 - \ln 2\right] - \left[3 \ln \tfrac{1}{2} - \ln \tfrac{5}{4}\right] \qquad \textbf{1}$$

$$= \ln\left(\tfrac{5}{4} \times \tfrac{8}{1} \times \tfrac{1}{2}\right)$$

$$= \ln 5 \approx 1.609 \qquad \textbf{1}$$

B6. (a) *Method 1 – separating the variables*

$$\sin x \frac{dy}{dx} - 2y \cos x = 0$$

$$\frac{dy}{dx} = 2\frac{\cos x}{\sin x} y$$

$$\int \frac{dy}{y} = 2 \int \frac{\cos x}{\sin x} dx \qquad \textbf{M1}$$

$$\ln y = 2 \ln(\sin x) + C \qquad \textbf{1}$$

$$= \ln\left(\sin^2 x\right) + C \qquad \textbf{1}$$

$$y = \exp\left(C + \ln(\sin^2 x)\right) \qquad \textbf{1}$$

$$= e^C \sin^2 x$$

Method 2 – using an integrating factor

$$\sin x \frac{dy}{dx} - 2y \cos x = 0$$

$$\frac{dy}{dx} - 2\frac{\cos x}{\sin x} y = 0 \qquad \textbf{1}$$

$$\int -2\frac{\cos x}{\sin x} dx = -2 \ln(\sin x) = \ln\left(\sin^{-2} x\right) \qquad \textbf{1}$$

Integrating factor = $\exp\left[\ln\left(\sin^{-2} x\right)\right] = \sin^{-2} x$

$$\frac{1}{\sin^2 x}\frac{dy}{dx} + \frac{-2 \cos x}{\sin^3 x} y = 0 \qquad \textbf{1}$$

$$\frac{d}{dx}\left(\frac{y}{\sin^2 x}\right) = 0$$

$$y = A \sin^2 x \qquad \textbf{1}$$

(b)

$$\sin x \frac{dy}{dx} - 2y \cos x = 3 \sin^3 x$$

$$\frac{dy}{dx} - 2\frac{\cos x}{\sin x} y = 3 \sin^2 x. \qquad \textbf{1}$$

Integrating factor is

$$\exp\left(\int -2\frac{\cos x}{\sin x} dx\right) = \exp\left(-2 \ln(\sin x)\right) = \frac{1}{\sin^2 x} \qquad \textbf{M1,1}$$

$$\frac{1}{\sin^2 x}\frac{dy}{dx} + \frac{-2 \cos x}{\sin^3 x} y = 3 \qquad \textbf{1}$$

$$\frac{d}{dx}\left(\frac{y}{\sin^2 x}\right) = 3$$

$$\frac{y}{\sin^2 x} = 3x + D$$

$$y = (3x + D) \sin^2 x \qquad \textbf{1}$$

[END OF MARKING INSTRUCTIONS]